DEVELOPMENTAL BIOLOGY 173, 382–395 (1996)
ARTICLE NO. 0034

REVIEW

Hox Genes and Chordate Evolution

Peter W. H. Holland* and Jordi Garcia-Fernàndez‡

**School of Animal and Microbial Sciences, The University of Reading, Whiteknights, Reading, RG6 6AJ, United Kingdom; and ‡Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Avenida Diagonal 645, 08071 Barcelona, Spain*

Hox genes are implicated in the control of axial patterning during embryonic development of many, perhaps all, animals. Here we review recent data on *Hox* gene diversity, genomic organization, and embryonic expression in chordates (including tunicates, amphioxus, hagfish, lampreys, teleosts) plus their putative sister group, the hemichordates. We consider the potential of comparative *Hox* gene data to resolve some outstanding controversies in chordate phylogeny. The use of *Hox* gene expression patterns to identify homologies between body plans both within the vertebrates and between the chordate subphyla is also discussed. Homology between the vertebrate hindbrain and an extensive region of amphioxus neural tube is suggested by comparison of *Hox-3* homologues and strengthened by new data on amphioxus *Hox-1* gene expression reported here. Finally, we give two examples of how *Hox* genes are giving glimpses into chordate developmental evolution. The first relates changes in *Hox* gene expression to transposition of vertebral identities; the second describes a correlation between vertebrate origins and *Hox* gene cluster duplication. We suggest that the simultaneous duplication of many classes of genes, often interacting in gene networks, allowed the elaboration of new developmental control mechanisms at vertebrate origins.

INTRODUCTION

The discovery of vertebrate *Hox* genes related to the homeotic selector (HOM) genes of *Drosophila* caused excitement in developmental biology, raising hopes for finding common patterning mechanisms in diverse animal species (e.g., McGinnis *et al.*, 1984; Carrasco *et al.*, 1984). *Hox* genes encode helix-turn-helix transcription factors and comprise one category of homeobox genes; the latter constitute a multigene family that probably arose by duplication and divergence from a single progenitor gene in early eukaryotes (for a comprehensive survey, see Duboule, 1994). The initial excitement intensified as it was realized that not only are *Hox* and HOM genes homologous (genealogically related) between phyla, but so is their arrangement into gene clusters: the common ancestor of chordates, arthropods, and nematodes must have possessed a *Hox* gene cluster (Duboule and Dollé, 1989; Graham *et al.*, 1989; Kappen *et al.*, 1989; Kenyon and Wang, 1991). The 38 *Hox* genes known in the mouse genome are arranged in four gene clusters, while the 8 *Drosophila* HOM genes (plus 4 other homeobox genes) are grouped into the Bithorax and Antennapedia complexes; each arrangement is descendent from a single cluster of at least 5 (probably more) *Hox* genes in the common ancestor of arthropods and chordates. It now seems likely that a linked cluster of *Hox* genes is a character shared by all metazoa and may be fundamental to controlling axial patterning in animals (Slack *et al.*, 1993).

The organizational and functional similarities between *Hox* genes from divergent animals have often been stressed and are certainly remarkable (reviewed by McGinnis and Krumlauf, 1992). But *Hox* genes hold evolutionary clues beyond simply being pointers to conservation of developmental mechanisms. In this review, we examine how analysis of *Hox* genes from different species is furthering our understanding of evolution, focusing specifically on our own phylum: the Chordata. Three subjects considered are (i) chordate phylogeny, (ii) homology between body plans, and (iii) the evolution of developmental mechanisms. To date, comparisons of *Hox* gene sequences and organisation have given limited insights into chordate phylogeny (although there is considerable potential). Descriptions of *Hox* gene expression in different species have contributed significantly to the recognition of homologies between divergent body plans (and hence patterns of morphological change). The third goal, understanding developmental

0012-1606/96 $12.00

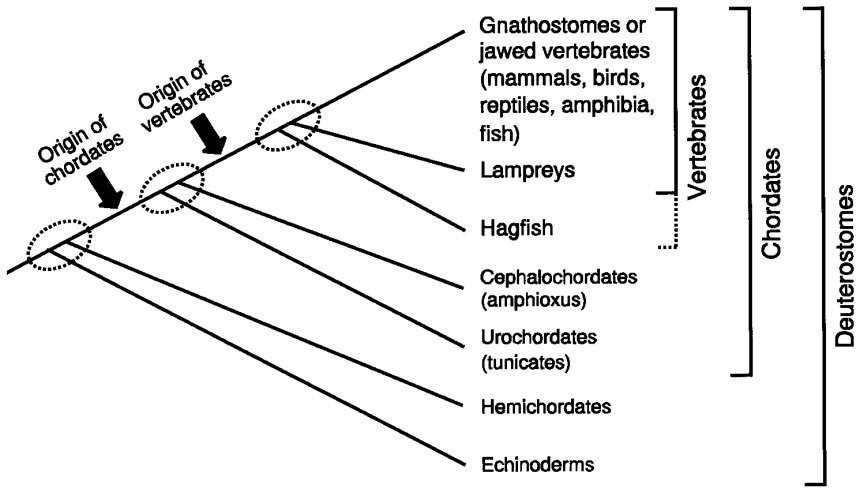

FIG. 1. Possible phylogenetic relationships between chordate taxa and the two other major deuterostome phyla, hemichordates and echinoderms. Dotted circles indicate principal regions of phylogenetic uncertainty, where alternative branching patterns have been postulated. For reasons discussed in the text, we include hagfish within the vertebrates; an alternative view considers gnathostomes and lampreys to be the only living taxa within the vertebrates (compare extent of solid and broken brackets).

mechanisms, is more ambitious; even so, *Hox* gene analyses have started to give glimpses into genetic constraints to developmental evolution and the coevolution of genes and development.

Before we discuss each of these subjects, we review current data on *Hox* gene presence, genomic organization, and expression patterns in the phylum Chordata and the closely related phylum Hemichordata. Within the Chordata, we consider data from the three major divisions or subphyla: the Urochordata (tunicates), the Cephalochordata (amphioxus), and the Vertebrata (including the hagfish, lampreys, teleost fish, and others). To put the comparative data in context, the putative phylogenetic relationships between these taxa discussed are depicted in Fig. 1; nodes of particular uncertainty in the phylogeny are highlighted. When genes can be compared clearly to *Hox* genes of mammals, we use the nomenclature of Scott (1992); these names indicate both the *Hox* gene cluster (a–d) and the ''paralogous group'' (1–13). Genes within a paralogous group are related by gene cluster duplication; for example, *Hoxa-1, Hoxb-1,* and *Hoxd-1* are group 1 genes located at the 3′ end of the *Hoxa, Hoxb,* and *Hoxd* clusters. When direct comparison to mammalian paralogous groups is less clear, we employ the broader groupings of Schubert *et al.* (1993). These divide *Hox* genes into ''anterior'' group genes (mammalian paralogous groups 1 to 3; *Drosophila lab, pb*), ''medial'' group genes (mammalian groups 4 to 8; *Dfd, Scr, Antp, Ubx, abd-A*), and ''posterior'' group genes (mammalian groups 9 to 13; *Abd-B*).

HEMICHORDATE *HOX* GENES

The phylum Hemichordata contains two principal groups of animals with rather dissimilar adult anatomy: the motile, worm-like enteropneusts (acorn worms) and the sessile, filter-feeding pterobranchs. Several features of their early embryology are shared with chordates, such as the radial pattern of early cleavage and posterior position of the blastopore; however, later development shows few similarities. Early claims that hemichordates have a homologue of the notochord seem unfounded, and although part of the enteropneust nervous system consists of a dorsal (and sometimes hollow) ''neurocord'' formed by ectodermal invagination, homology with the chordate neural tube is unproven (Brusca and Brusca, 1990).

Pendleton *et al.* (1993) included the enteropneust *Saccoglossus kowalevskii* in an extensive phylogenetic survey for *Hox* genes. Using degenerate PCR primers (designed to amplify genes related to mammalian paralogous groups 1 to 10), Pendleton *et al.* cloned short fragments from nine *Hox* genes. Partial homeobox sequences are often insufficient to relate *Hox* genes to precise paralogous groups; even so, the PCR data in this case are sufficient to draw several conclusions. *Saccoglossus* has retained representatives of the anterior, medial, and posterior groups of *Hox* gene, thought to have originated very early in animal evolution (Schubert *et al.*, 1993). Three *Saccoglossus* PCR clones in the anterior group can each be putatively homologized to mammalian *Hox* paralogous groups 1, 2, and 3. Five of the *Saccoglossus* clones derive from medial group genes and one from a posterior group gene (related to *Abd-B*). The precise evolutionary relationships between the medial group *Hox* genes of hemichordates and other taxa cannot be deduced from currently available DNA sequence; considering the uncertainty over hemichordate affinities and monophyly (Fig. 1; see also Peterson, 1995), resolution of these gene relationships could have considerable phylogenetic importance. There is certainly a need for more extensive *Hox* gene DNA sequence

data, plus gene linkage analysis, in enteropneusts, pterobranchs, and echinoderms.

UROCHORDATE *HOX* GENES

The urochordates (tunicates) comprise a subphylum within the Chordata and include the sessile ascidians plus the pelagic salps and larvaceans. The ascidians have proved popular and informative model systems in developmental biology (Satoh, 1994), since they combine advantages offered by a defined cell lineage with the development of a chordate body plan (including notochord, dorsal nerve cord, and lateral muscle cells). Their phylogenetic position also makes them of crucial importance to understanding chordate evolution and vertebrate origins.

The cloning and analysis of urochordate *Hox* genes, however, has not proved straightforward. The first ascidian homeobox gene reported was a non-*Hox* cluster gene, *AHox1,* from *Halocynthia roretzi* (Saiga *et al.,* 1991). It is surprising that *AHox1,* a homologue of the divergent *Drosophila H2.0* gene, was the most strongly hybridizing gene detected by Southern hybridization using a *Hox* gene probe. In a subsequent search for ascidian *Hox* genes, degenerate PCR was applied to four species of ascidian (Ruddle *et al.,* 1994). Single *Hox* gene sequences were cloned from three species (*CI-1* from *Ciona intestinalis*, *SC-6* from *Styela clava*, *MO-4* from *Molgula citrina*); surprisingly, no ascidian yielded more than a single gene. A single *Hox* gene was also identified by PCR on genomic DNA from a larvacean, *Oikopleura dioica* (Holland *et al.,* 1994b).

Several hypotheses were proposed to explain these puzzling results (Saiga *et al.,* 1991; Holland *et al.,* 1994b; Ruddle *et al.,* 1994) including extreme divergence of some urochordate *Hox* gene sequences or insertion of introns within some of the homeoboxes. Either situation could disrupt probe and primer hybridization or reduce efficiency of PCR amplification. A more radical suggestion is that few *Hox* genes are present in urochordate genomes; this would imply gene loss, because phylogenetic considerations argue that urochordates evolved from animals possessing a *Hox* gene cluster. Difficult to reconcile with the gene loss hypothesis is the fact that the genes amplified by PCR from different species do not seem to be orthologues (Ruddle *et al.,* 1994); this would demand independent loss in multiple lineages.

The explanation is now emerging as more extensive *Hox* gene sequences are obtained from ascidian genomic and cDNA libraries. Ge *et al.* (1994) describe isolation of a *Styela clava Hox* cDNA, designated *AHox2,* plus a genomic clone of its putative *Styela plicata* orthologue, *AHox3.* The *S. clava AHox2* sequence (Ge *et al.,* 1994) differs from the *S. clava SC-6 Hox* PCR clone (Ruddle *et al.,* 1994), implying that this ascidian does possess more than one *Hox* gene. The *AHox2/AHox3* homeodomain is unusually divergent from other chordate *Hox* gene sequences, having residues characteristic of both the medial and posterior group genes.

Unusual sequence divergence of ascidian homeodomains has also been found by Di Gregorio *et al.* (1995) in an extensive survey of homeobox genes from *C. intestinalis.* These authors describe five *Ciona Hox* genes (plus nine other homeoboxes) isolated by genomic library screening using a degenerate oligonucleotide. The *Hox* genes include representatives of the anterior, medial, and posterior groups, but sequence divergence makes direct comparisons to mammalian paralogous groups tentative. The data reveal another peculiarity of *Ciona Hox* genes that helps explain low efficiency of PCR cloning: the presence of introns within some homeoboxes (spanning 2.1 kb for the putative group 3 homologue *CiHbox1;* Di Gregorio *et al.,* 1995). It is interesting that three of the *Ciona Hox* genes cloned belong to the posterior group (of which *Drosophila* has only *Abd-B*, while mammals have five paralogous groups); this suggests that duplication of posterior group genes occurred early in chordate evolution. It is not known if this expansion of posterior group *Hox* genes occurred on the chordate stem lineage or on multiple occasions during chordate evolution. At least two of the *Ciona* posterior group *Hox* genes are reported to be physically linked, confirming cluster organization (Di Gregorio *et al.,* 1995).

Similar conclusions of multiple ascidian *Hox* genes, gene cluster organization, and intron presence have been obtained for another ascidian species, *H. roretzi* (H. Saiga, personal communication; Katsuyama *et al.,* 1995). These authors identified five *Halocynthia Hox* genes using PCR: a group 1 gene, a group 2 gene, two genes from the medial group, and one from the posterior group. Screening of genomic libraries has shown at least three of these genes are linked (Katsuyama *et al.,* 1995). A cDNA clone was also isolated for the group 1 homologue, *HrHox-1,* and expression in ascidian embryos and larvae analyzed. *HrHox-1* is expressed in epidermal cells and in the developing central nervous system, in a spatially restricted domain along the anteroposterior axis (Katsuyama *et al.,* 1995). The epidermal expression (in a diffuse region excluding the anterior and posterior termini) is first detectable in early neurulae, although the CNS expression is not seen until the tailbud stage. The latter expression can be resolved into two patches: an anterior band at the junction between the cerebral vesicle and neural tube (that persists to the swimming tadpole larval stage) and a patch marking the presumptive neural tube (this staining does not persist). Very weak *HrHox-1* expression was detected at two other sites in the CNS; interestingly, both lie anterior to the principal expression domain and are associated with the principal sense organs of the larva (the ocellus and otolith).

Katsuyama *et al.* (1995) show that retinoic acid greatly up-regulates the expression of *HrHox-1* and broadens its spatial domain rostrally; associated phenotypic alterations at the anterior end of the tadpole larva are also quite dramatic. Retinoic acid has a less marked effect on an expression of an ascidian posterior group gene, *HrHox-10;* this result is consistent with colinearity in retinoic acid responsiveness of ascidian *Hox* genes, as is well characterized for

their vertebrate homologues (e.g., Simeone *et al.,* 1990; Papalopulu *et al.,* 1991).

It seems certain that continued study of the unusual *Hox* genes of ascidia will give further insights into the molecular evolution of *Hox* genes and possible constraints to their sequence divergence or genomic arrangement. Distinguishing between selectively constrained and neutral features of *Hox* gene organization should be possible by comparison between different tunicates; *Hox* genes are being studied in *Ciona* (Di Gregorio *et al.,* 1995), *Styela* (Ge *et al.,* 1994), *Halocynthia* (Katsuyama *et al.,* 1995), and *Phallusia* (W. J. Gehring and P. Baumgartner, personal communication). The deployment of *Hox* genes in relation to a fixed cell lineage is also of great interest, and further studies complementary to those discussed above will prove an interesting contrast to the other chordates. A rich legacy of experimental embryology, together with techniques for embryo and gene manipulation (Satoh, 1994), should facilitate examination of the regulation and role of ascidian *Hox* genes.

AMPHIOXUS *HOX* GENE ORGANIZATION

Amphioxus is a common name given to animals of the genus *Branchiostoma:* one of two genera currently recognized within the subphylum Cephalochordata or Acrania. Amphioxus has been popular with embryologists and evolutionary biologists for over a century (for review see Willey, 1894) and still features in more enlightened embryology textbooks and courses. Historically, the animal's popularity reflected its generalized chordate anatomy. For example, amphioxus has a dorsal hollow nerve cord, a prominent axial notochord, and a metameric series of bilaterally paired segmented muscle blocks (developed from somites) along the entire anteroposterior axis. The simple arrangement or ''schematic clearness'' of the body organization led Willey (1894) to assert it is ''no wonder that Amphioxus should serve as a refuge to the perplexed embryologist.''

The traditional view that amphioxus anatomy reflects that of vertebrate ancestors is still popular, although certainly not universally held (Jefferies, 1986). That cephalochordates are the sister group of vertebrates is still the most widely held view and one that has gained recent support from phylogenetic analyses of 18S ribosomal DNA sequences (Wada and Satoh, 1994) and cladistic analysis of morphological characters (Peterson, 1995). Despite the similarities, there are also substantial differences between amphioxus and vertebrate anatomy; for example, amphioxus lacks cranial ganglia, paired cranial sense organs, and an elaborate craniofacial region dominated by neural crest-derived mesenchyme. These differences suggest that a suite of innovations evolved on the vertebrate lineage, after the cephalochordates diverged; innovations that required the evolution of new developmental programs (Gans and Northcutt, 1983; Holland, 1988, 1992).

It is no surprise that an animal occupying such a key phylogenetic position has received the attention of several laboratories interested in *Hox* gene evolution. The first amphioxus homeobox sequence, *AH1,* was cloned soon after the discovery of the homeobox (W. McGinnis, unpublished data); recent sequence comparisons show this to be *AmphiHox-7,* the amphioxus homologue of mammalian paralogous group 7 genes. Holland *et al.* (1992) reported the cloning and expression pattern of an amphioxus *Hox* gene encoding a protein with extensive stretches of sequence similarity to the mouse Hoxb-3 protein. This gene is clearly a homologue of vertebrate *Hox* paralogous group 3 genes and hence is designated *AmphiHox-3.*

The number of *Hox* genes in the amphioxus genome was addressed by two laboratories using degenerate PCR, but with differing conclusions. Amphioxus was included in the broad phylogenetic survey of *Hox* genes performed by Pendleton *et al.* (1993), who suggested there were probably two *Hox* gene clusters in the amphioxus genome; Holland *et al.* (1994a) suggested that there was one. The contradiction was subsequently resolved by exhaustive genomic and cDNA screening that conclusively demonstrated the presence of a single *Hox* gene cluster in amphioxus (Garcia-Fernàndez and Holland, 1994). The reasons for uncertainty and inconsistency from the initial PCR surveys are worth considering, since they offer useful lessons to other studies. One problem facing interpretation of PCR clones is the difficulty (in some cases impossibility) of relating short PCR sequences to particular paralogous groups of mammalian *Hox* genes. Added to this is the problem of how to distinguish genes (loci) from alleles at the same locus. In amphioxus, this proved a particular problem for the *AmphiHox-2* gene; subsequent cloning from genomic libraries revealed sequence polymorphism in this gene. Finally, analysis using gene-specific primers (Garcia-Fernàndez and Holland, 1994) found that two clones reported by Pendleton *et al.* (1993) were not derived from the amphioxus genome (nor from recognizable laboratory contaminants); a reminder of the problems caused by extreme sensitivity of PCR, despite careful precautions being taken.

The organization of the single amphioxus *Hox* gene cluster, as determined by genomic walking, transpired to be of particular interest (Garcia-Fernàndez and Holland, 1994). A 270-kb genomic region contains 10 *Hox* genes (*AmphiHox-1* to *-10*); on the basis of sequence, it was concluded that each gene is a homologue of a different paralogous group of mammalian *Hox* genes (groups 1 to 10, respectively). The physical gene order in amphioxus is colinear with the assignment number; *AmphiHox-1* being the most 3′ gene in the cluster (Fig. 2). The published study did not address whether genes homologous to the extreme 5′ paralogous groups were present in amphioxus (mammals possess group 11 to 13 genes in this region). Subsequent genomic walking in amphioxus has demonstrated the presence of at least two additional *Hox* genes 5′ of *AmphiHox-10* (J.G.F., P.W.H.H., and S. A. J. Thompson, unpublished data); the evolutionary relationships between these genes and mammalian posterior paralogous groups are currently under investigation.

The results of the amphioxus genomic walk demonstrate

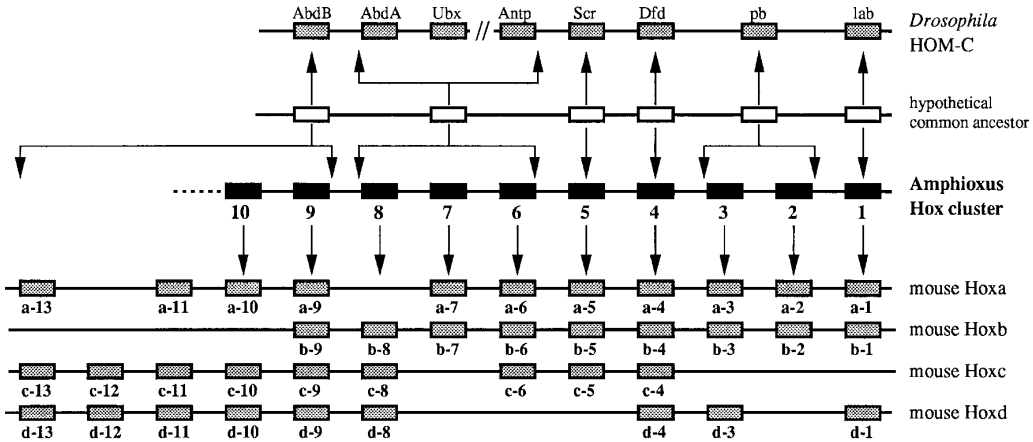

FIG. 2. Suggested evolutionary relationship between *Drosophila*, mammalian, and amphioxus *Hox* gene clusters, together with that inferred for the hypothesized common ancestor of chordates and arthropods. The amphioxus *Hox* gene cluster has an archetypal organization in relation to the vertebrate clusters. Reproduced from Garcia-Fernàndez and Holland (1994), with permission from *Nature*, copyright 1994, Macmillan Magazines Ltd.

that duplication of an ancestral *Hox* gene cluster (to give four) occurred on the lineage leading to higher vertebrates, subsequent to divergence of the cephalochordate lineage. They are also consistent with the ''two phase'' model for vertebrate *Hox* gene evolution proposed previously (Kappen *et al.*, 1989; Kappen and Ruddle, 1993); this stated that duplication of the *Hox* gene cluster occurred after it had been substantially expanded by tandem gene duplications. One very surprising conclusion from the amphioxus genomic walk is that there must have been no tandem *Hox* gene duplications along the evolutionary lineage leading to amphioxus, at least within the anterior and medial groups of *Hox* genes. Whether this conclusion extends to the most 5′ *Hox* genes (the posterior group) is uncertain at present. The overall picture, therefore, is that amphioxus has retained an archetypal *Hox* gene cluster organization, remarkably similar to that inferred for a direct ancestor of the vertebrates. This conservation of developmental patterning genes, despite 520 million years of independent evolution from the chordate stem linage, parallels the long-standing view of amphioxus as an archetypal chordate, retaining key anatomical characters of the prevertebrate body plan. Amphioxus may be uniquely suited for examining the regulation of chordate *Hox* genes within a single gene cluster.

AMPHIOXUS *HOX* GENE EXPRESSION

The expression of the *AmphiHox-3* gene in amphioxus embryos and larvae has been examined using *in situ* hybridization (Holland *et al.*, 1992, 1994a,b); an investigation facilitated by the development of reliable *in vitro* spawning methods for *Branchiostoma floridae* by N. D. Holland and L. Z. Holland (1989, 1993). The *AmphiHox3* gene is expressed in cells of the neural furrow of neurula stage amphioxus embryos and persists in the nerve cord through embryonic and larval development, to juvenile and adult stages. Expression is not seen in the most rostral regions of the presumptive nerve cord and respects a temporally stable rostral limit around the position of the somite four/five boundary (Fig. 3). Restriction to a particular region of the body axis is a character shared with vertebrate *Hox* genes; however, in the case of *AmphiHox-3*, this region-specific expression is seen only in presumptive nerve cord. Although the gene is also expressed in a patch of mesodermal cells, this is always posterior and cannot be considered to respect an axial position.

Holland *et al.* (1992, 1994a,b) make two evolutionary inferences from the *AmphiHox-3* expression pattern. First, that the role of chordate *Hox* genes in region-specific patterning of the nerve cord is more ancient than a similar role in the mesoderm. Second, that the vertebrate hindbrain is homologous to an extensive region of amphioxus nerve cord. The latter inference uses the anterior limit of *Hox* genes as a clue to homology (see later). Each inference could be strengthened or weakened as other amphioxus *Hox* genes are examined. In Fig. 3, we compare the expression pattern of *AmphiHox-3* with that of the most 3′ amphioxus *Hox* gene, *AmphiHox-1*. The homologue of mammalian paralogous group 1 is of particular interest, in view of the deviation from spatial colinearity shown by mouse or chick *Hoxa-1* and *Hoxb-1* (each has a rostral expression limit posterior to that of the group 2 genes; Wilkinson *et al.*, 1989; Murphy and Hill, 1991; Prince and Lumsden, 1994).

AmphiHox-1 RNA, like *AmphiHox-3*, is detected in both neurectoderm and mesoderm of amphioxus embryos and larvae; *AmphiHox-1* is also expressed in some epidermal cells (Figs. 3C and 3D). The expression pattern of *Amphi-*

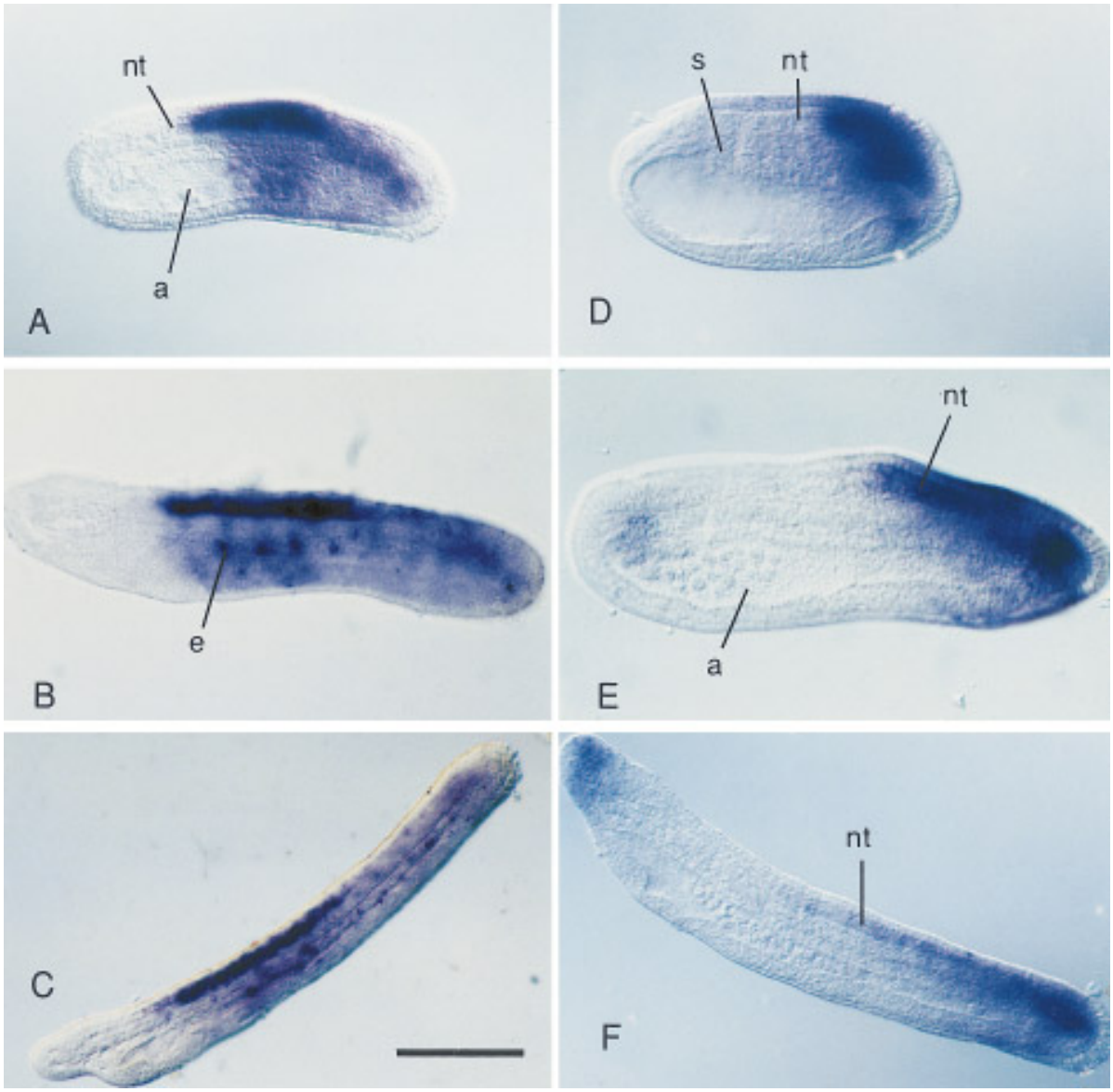

FIG. 3. Expression of *AmphiHox-1* (A–C) and *AmphiHox-3* (D–F) in amphioxus embryos and larvae, revealed by nonradioactive whole mount *in situ* hybridization. (A–C) *AmphiHox-1* expression at 16 hr (late neurula), 20 hr and 36 hr (larva), respectively; (D–F) *AmphiHox-3* expression at 13 hr (neurula), 18 hr and 30 hr, respectively. Anterior is to the left in all figures. Scale bar, 100 μm in A–C or 80 μm in D–F. The strongest *AmphiHox-1* expression is within the developing neural tube, with clear anterior and posterior limits. Weaker expression is seen in other tissues, including some epidermal cells; this also respects an anterior boundary. *AmphiHox-3* expression within the developing neural tube has a clear anterior boundary, caudal to that of *AmphiHox-1*. a, archenteron; e, epidermal expression; nt, presumptive neural tube; s, first somite.

Hox-1 in neurectoderm is intriguing; unlike *AmphiHox-3* expression, the strongest patch of staining has both rostral and caudal limits at the neurula stage. This is reminiscent of one its homologues, mouse *Hoxb-1*, which is predominantly restricted to rhombomere 4 by 8.5 days post coitum (Wilkinson *et al.*, 1989; Murphy and Hill, 1991). As amphioxus development proceeds, the rostral limit to *AmphiHox-1* expression remains stable (approximately the somite 3/4 level). Much lower levels of *AmphiHox-1* expression are detected in a domain encompassing the embryo, but again respecting a rostral boundary; a bilateral series of epidermal cells expresses *AmphiHox-1* at a much higher level than the surrounding epidermal cells (Fig. 3).

The *AmphiHox-1* and *AmphiHox-3* expression patterns

are both consistent with roles in region-specific patterning of the neural tube (*AmphiHox-1* may also play such a role in some epidermal cells).

HAGFISH AND LAMPREY *HOX* GENES

Of the many lineages of jawless fish that originated early in vertebrate evolution, only two have living representatives: the myxinoids (hagfish) and petromyzonts (lampreys). Both taxa share with the higher vertebrates a notochord, dorsal neural tube, tripartite brain, neural crest-derived tissues, cranial ganglia, segmented muscle blocks, and the paired sense organs; they lack jaws and paired fins. Anatomical, physiological, and paleontological evidence suggests that hagfish are excluded from a clade containing lampreys and jawed vertebrates (Fig. 1); in contrast, 18S rDNA sequence analysis suggests that hagfish and lampreys form a natural group (Stock and Whitt, 1992). Because of this phylogenetic uncertainty, here we include the hagfish within the taxon Vertebrata (in contrast to Janvier, 1993).

The extreme difficulty of obtaining hagfish embryos greatly restricts the scope for molecular developmental studies in these animals. Nonetheless, the group occupies such an important evolutionary position that cloning and analysis of *Hox* genes is certainly warranted. The *Hox* gene complement is currently being surveyed in the hagfish *Eptatretus stoutii* (G. Wagner and W. Bailey, personal communication). Preliminary data suggest the presence of up to four *Hox* gene clusters; whether these are resultant from the same gene cluster duplication as the mammalian clusters in unknown.

A degenerate PCR screen for *Hox* genes has also been performed on the sea lamprey *Petromyzon marinus* (Pendleton *et al.*, 1993). Nineteen different *Hox* genes were distinguished, including two genes from paralogous group 1 and five assignable to groups 9 or 10 (posterior group). Three of the genes identified by PCR were also found to be linked on a single cosmid clone, confirming that lamprey *Hox* genes are clustered. Pendleton *et al.* conclude that the *P. marinus* probably has three *Hox* gene clusters; however, a two or four cluster state cannot be excluded. Preliminary PCR surveys in the river lamprey *Lampetra fluviatilis* again suggest multiple *Hox* gene clusters (A. C. Sharman and P.W.H.H, unpublished data). Also of interest is the isolation of a *L. fluviatilis* clone assignable to paralogous group 13 of mammals, confirming the early expansion of the posterior group genes.

The hagfish and lamprey results are important, since they indicate that duplication of a single *Hox* gene cluster occurred close to the origin of the vertebrates. The implications for the mechanisms of developmental evolution are discussed later. Determination of more extensive lamprey and hagfish *Hox* gene sequences, plus precise elucidation of cluster number, should give insights into the order of cluster duplication (were there intermediate states of two or three clusters?), whether or not cluster duplication occurred independently in different vertebrate lineages (hagfish, lamprey, jawed vertebrate), and the mutational basis for duplication (by tetraploidy or local duplication?). It should also help resolve the phylogenetic relationships between hagfish, lampreys, and jawed vertebrates.

HOX GENES OF TELEOST FISH

It is well known that human and mouse genomes each possess four *Hox* gene clusters, both species probably possessing the same 38 *Hox* genes (McGinnis and Krumlauf, 1992). From an evolutionary perspective, we wish to know how ancient the cluster duplications are, and how the 38 genes arose. Among the many lineages of jawed vertebrate, teleost fish are currently proving particularly informative in answering these questions. *Hox* genes have been cloned from several teleosts including zebrafish, goldfish, salmon, striped bass, puffer fish, and killifish (e.g., Kappen *et al.*, 1993; Aparicio *et al.*, 1995; Misof and Wagner, 1995; Pavell and Stellwag, 1995). Developmental expression has been analyzed for several genes, particularly in the zebrafish, *Brachydanio rerio* (e.g., Molven *et al.*, 1992).

Only recently has the genomic organization of teleost *Hox* genes been mapped in detail. Current data suggest the presence of four *Hox* gene clusters in the genomes of the zebrafish (Sordino *et al.*, 1995; D. Duboule, personal communication) and the puffer fish, *Fugu rubripes* (S. Aparicio and S. Brenner, personal communication); these probably have a 1:1 correspondence with the four mammalian *Hox* gene clusters.

An intriguing property of the mammalian *Hox* gene clusters is that no cluster contains representatives of all 13 paralogous groups. The implication is that after duplication of the ancestral chordate *Hox* gene cluster, some genes were deleted or scrambled within each cluster. It is interesting to determine if zebrafish and puffer fish have precisely the same *Hox* gene complement as mammals. The possibility that differences may exist was raised by Molven *et al.* (1992), who described a zebrafish *Hox* gene that could be assigned to paralogous group 3, but was not a clear orthologue of any mammalian group 3 gene. A recent PCR screen for *Hox* genes in the killifish, *Fundulus heteroclitus*, is informative in this context (Misof and Wagner, 1995). Partial sequences from 27 *Hox* genes were recovered, which were compatible with (but not proof of) four gene clusters. Tentative assignments to mammalian paralogous groups, plus statistical considerations, suggest that the killifish and mammalian *Hox* clusters have lost different genes. For example, 4 distinct killifish sequences can be assigned to paralogous group 1, whereas mammals possess 3 such genes (*Hoxa-1, Hoxb-1,* and *Hoxd-1*).

The teleost data suggest that the common ancestor of teleost fish and mammals possessed four *Hox* gene clusters and that these were retained in each lineage. The precise gene complement of each cluster was probably not stable during vertebrate radiation, however, and different lineages

lost different genes. Whether the resultant differences are functionally relevant, and whether they will be informative as indicators of phylogenetic affinity, remains to be investigated.

HOX GENES AND CHORDATE PHYLOGENY

Chordate phylogeny is a popular subject for debate and disagreement and one with relevance to developmental biologists. Unless phylogenetic relationships between living organisms are known, comparative developmental biology can give limited insight into the evolution of developmental mechanisms. The phylogeny proposed in Fig. 1 is a popular hypothesis of the relationships between chordates and their close relatives, but alternative schemes have been proposed and may have different implications for embryological evolution (Holland and Graham, 1995). One use of an accurate phylogeny is that it helps distinguish between taxa that have always lacked a particular (developmental) character and taxa that have secondarily lost the character. For example, if tunicates (not amphioxus) are the sister group of vertebrates, this would suggest tunicates have secondarily lost overt segmentation, pushing the origin of segmentation back to chordate origins. An implication of this unorthodox phylogeny is that comparison between tunicates, amphioxus, and vertebrates would be of little use in understanding the origin of segmentation. Similarly, a sister group relationship between echinoderms and chordates (to the exclusion of hemichordates) would be consistent with a controversial scenario placing the origins of neural crest cells, paired sense organs, and the notochord very early in deuterostome radiation (Jefferies, 1986); if this scenario proves correct, comparison between amphioxus and vertebrates would be of little use in studying the origin of neural crest cells.

Molecular analyses of chordate phylogeny have generally used sequence comparisons between ribosomal RNA genes; these support some aspects of Fig. 1, but certainly not all the details (Wada and Satoh, 1994; Halanych, 1995). Although ribosomal RNA is a widely used and powerful phylogenetic indicator, it does have limitations; for example, it may be subject to a degree of convergent evolution, to differing substitution rates between lineages, and to saturation of mutations at variable sites (Philippe *et al.*, 1994). These do not often cause serious problems, but they create difficulties for resolving nodes spaced closely in time (such as encountered in early chordate evolution). The same limitations complicate all attempts to use primary sequence data to investigate ancient phylogenetic events. There is growing interest, therefore, in finding alternative strategies for molecular phylogenetic inference; ideally, these would be based on types of mutation that are sufficiently rare and complex to be immune from the problems of convergent evolution or reversion. Possible characters include mitochondrial gene order, variants in the mitochondrial genetic code, retroposon integrations, gene losses, gene duplications, and, with some caveats, intron positions (Kido *et al.*, 1991; Sankoff *et al.*, 1992; Philippe *et al.*, 1994; Ruddle *et al.*, 1994; Holland and Graham, 1995).

It has been suggested that *Hox* gene clusters may hold useful phylogenetic information of this type (Philippe *et al.*, 1994). To date, insufficient comparative data have been accumulated to be particularly useful, but the early indications are promising (G. Balavoine and A. Adoutte, personal communication). For example, additional introns seem to have been inserted into some tunicate *Hox* genes during evolution (Di Gregorio *et al.*, 1995); these may prove useful characters for resolving tunicate phylogeny, if homology of intron insertion can be demonstrated between taxa. The number of *Hox* gene clusters may help resolve hagfish and lamprey affinities, if the multiple clusters are shown to be directly homologous between vertebrate lineages. The presence of a single and archetypal *Hox* gene cluster in amphioxus is sufficient to refute any suggestion that cephalochordates descended from vertebrates by secondary degeneration; although, to be fair, this idea was generally rejected by the turn of the present century (H. Gee, personal communication; see comments of Goodrich in Gaskell *et al.*, 1909). Finally, resolution of early vertebrate phylogeny could be assisted if *Hox* gene loss, subsequent to cluster duplication, indeed followed different paths in different vertebrate lineages (Misof and Wagner, 1995).

HOMOLOGY BETWEEN BODY PLANS

The term homology is most widely used in its ''historical'' context, referring to the common evolutionary ancestry of a character (molecular, morphological, or other) within or between species (Roth, 1984; Reeck *et al.*, 1987; Minelli, 1993; Hall, 1994; Dickinson, 1995). When discussing homology of morphological characters, the definition may also have a developmental component. For example, Roth (1984) defined the basis of homology as ''the sharing of pathways of development, which are controlled by genealogically-related genes''; Minelli (1993) stresses the ''common informational background'' for the characters. Since *Hox* genes are clearly part of the ''informational background'' involved in the controlling ''pathways of development,'' it seems feasible to use them as indicators of morphological homology. The principal role of vertebrate *Hox* genes seems to be maintenance and interpretation of positional information along the anteroposterior body axis, with the rostral region of expression marking the most important site of action for each gene (McGinnis and Krumlauf, 1992). This interpretation of *Hox* gene function, together with the above view of homology, implies that homologous regions or structures in different species may be revealed by comparing the rostral expression limits of *Hox* genes.

A potential pitfall with this strategy, not obvious at first, relates to how far *Hox* genes lie along the ''pathway of devel-

opment'' leading to a particular phenotypic character (and hence how evolution acts upon such pathways). Although vertebrate *Hox* genes are involved in ensuring that particular characters develop at appropriate axial positions, this need not imply an unbreakable link between *Hox* genes and phenotype, as pointed out by Burke *et al.* (1995). One possibility is that vertebrate *Hox* genes mark fixed axial positions (e.g., particular somite numbers); they then control the development of whichever phenotypic character is appropriate to that position in any given species. Alternatively, the genes could lie further along the developmental pathway and mark presumptive phenotypic characters; in this case their expression patterns would be expected to shift axial level between divergent vertebrate taxa (Burke *et al.*, 1995). If we are to use *Hox* gene expression patterns as indicators of homology, it is desirable that they act in accordance with the second model. Homology as detected by *Hox* gene expression would then be comparable to homology as suggested by morphological similarity.

Burke *et al.* (1995) undertook an extensive comparison of *Hox* gene expression between mouse and chick embryos and found strong support for the second of the models described above. For example, *Hoxc-6* expression has an anterior expression limit at somite level 12–13 in mouse embryos, but at level 19–20 in the chick; these different axial levels correspond to equivalent phenotypes (the first thoracic vertebra). The data of Burke *et al.* (1995) reveal that despite evolutionary changes in somite number or axial formulae, *Hox* gene expression limits remain tightly correlated with particular phenotypic characters.

Hox gene expression limits, therefore, can be used to reveal homology between higher vertebrate body plans (such as *Hoxc-6* marking the first thoracic vertebra). This approach should also allow homology to be identified when histological examination has not revealed underlying developmental equivalence; a situation we refer to as ''cryptic homology.'' It is important to consider, however,the phylogenetic distance beyond which *Hox* gene expression patterns cease to be indicators of homology. This is not easily determined. Part of the problem is that homology is not an absolute property (Minelli, 1993); as developmental control pathways diverge between species over evolutionary time, so must the level of homology. We suggest that *Hox* genes will be very clear indicators of homology when the developmental roles of *Hox* genes are well conserved between taxa and when precisely orthologous *Hox* genes exist and can be recognized. Current data suggest these criteria hold for comparison between bony vertebrate lineages (e.g., chick vs mouse vs zebrafish) and perhaps also for comparison between lampreys, hagfish, and jawed vertebrates (if the multiple gene clusters are directly homologous).

What of more distant comparisons? For understanding the evolution of chordate body plans, it would be particularly useful to identify cryptic homologies between animals from the three chordate subphyla: amphioxus, tunicates, and vertebrates. A problem is that none of the *Hox* genes in amphioxus or tunicates are orthologues of single *Hox* genes in vertebrates, since gene cluster duplication occurred on the vertebrate lineage. We argue that it is still valid to use *Hox* genes as clues to morphological homology between taxa, but only when (a) single amphioxus or tunicate *Hox* genes can be assigned precisely to particular paralogous groups of vertebrate *Hox* genes and (b) the vertebrate *Hox* genes within a paralogous group share an axial limit to expression. The first criterion seems to hold for most, if not all, amphioxus *Hox* genes (Fig. 2; Garcia-Fernàndez and Holland, 1994); it remains to be seen if it also holds for tunicates. Current data suggest that the second criterion holds for vertebrate paralogous groups 1 and 3, possibly group 4, but not group 2 or some of the most posterior groups (Gaunt, 1991; Hunt *et al.*, 1991; Prince and Lumsden, 1994).

We suggest, therefore, that the amphioxus homologues of vertebrate *Hox* paralogous groups 1, 3, and 4 will hold clues to cryptic homology between the amphioxus and vertebrate body plans. This comparison can be drawn for axial positions in the neural tube, but not for the somitic mesoderm where amphioxus *Hox* genes have not yet been shown to respect axial limits. We described above the embryonic and larval expression patterns of two of these amphioxus genes, *AmphiHox-1* and *AmphiHox-3* (Fig. 3), assignable to vertebrate *Hox* paralogous groups 1 and 3 (Holland *et al.*, 1992; Garcia-Fernàndez and Holland, 1994). In mouse and chick embryos, genes from paralogous groups 1 and 3 are expressed in the developing neural tube, with rostral limits to expression located in the hindbrain. Group 1 genes (*Hoxa-1, Hoxb-1,* and *Hoxd-1*) have a limit at the rhombomere 3/4 boundary; for group 3 genes (*Hoxa-3, Hoxb-3,* and *Hoxd-3*), the limit is the rhombomere 4/5 boundary (Hunt *et al.*, 1991). Since the developing amphioxus neural tube does not show external segmentation, *Hox* gene expression must be described with reference to the adjacent somite pairs. *AmphiHox-3* has a rostral limit in the neural tube adjacent to the division between somite pairs 4 and 5; the limit for *AmphiHox-1* is approximately one somite more anterior.

In Fig. 4, we use these *Hox* gene expression limits as clues to homology between amphioxus and vertebrates body plans. The comparison helps resolve a long-standing uncertainty about homology between the amphioxus nerve cord and regions of the vertebrate brain. Willey (1894) contrasted several alternative schemes proposed in the last century, including Stieda's popular view that the entire vertebrate brain is homologous to the small ''cerebral vesicle'' of amphioxus (a dilation of the most rostral region of the nerve cord). The *Hox* gene expression data challenge this view, since they imply that the vertebrate hindbrain is homologous to an extensive region of the amphioxus neural tube, extending posterior to somite five. In this view, the cerebral vesicle can only be homologous to regions rostral to the hindbrain. This conclusion has gained recent support from detailed examination of the cellular architecture of the amphioxus cerebral vesicle, determined by serial transmission electron microcopy (Lacalli *et al.*, 1994). This study suggested that the amphioxus cerebral vesicle contains struc-

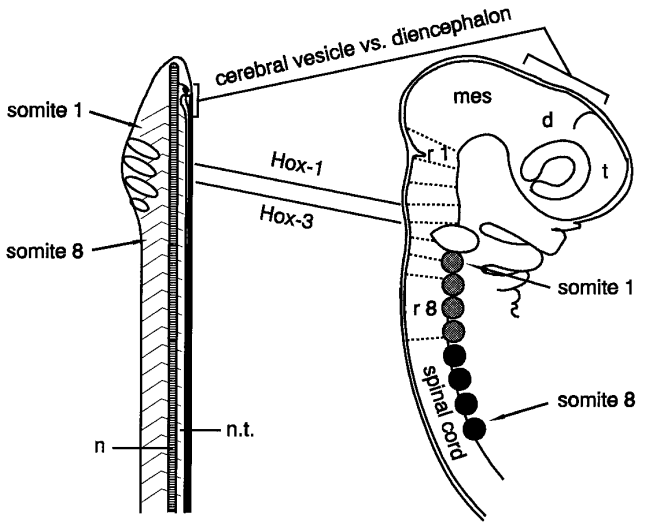

FIG. 4. Homologous regions between the neural tube of an amphioxus larva (left) and the brain of a generalized higher vertebrate embryo (right), as suggested by cellular organization (Lacalli *et al.*, 1994) and *Hox* gene expression (Holland *et al.*, 1992; plus this report). The rostral limit of *AmphiHox-1* gene expression in the amphioxus neural tube is compared with paralogous group 1 *Hox* genes of mouse and chick; *AmphiHox-3* is compared with paralogous group 3 *Hox* genes. d, diencephalon; mes, mesencephalon; n, notochord; n.t., neural tube; r, rhombomere; t, telencephalon.

tures and cell types homologous to those of the diencephalon; recent data also point to a short midbrain homologue posterior to the cerebral vesicle (T. C. Lacalli, personal communication).

These clues to homology have evolutionary implications, if we first assume the polarity of transformation between the types of body plan. Assuming that the common ancestor of cephalochordates and vertebrates possessed a body plan similar to that of modern amphioxus (at least in brain organization), then we conclude that the vertebrate hindbrain evolved by elaboration of an extensive region of neural tube. The diencephalon (and possibly other brain regions) evolved from an already specialized terminal region. This evolutionary elaboration of the brain from an extensive precursor probably occurred concomitantly with a major remodeling of the entire head region, proposed to have occurred during the origin of vertebrates (Gans and Northcutt, 1983; Holland, 1992).

EVOLUTIONARY MODIFICATION OF DEVELOPMENTAL MECHANISMS

It is not the intention of this article to consider the vast diversity of mechanisms by which development could be modified during evolution; broad perspectives can be found in Raff and Kaufman (1983), Hall (1992), and Akam *et al.* (1994). Here we ask whether mutations in the number, organisation or regulation of *Hox* genes were involved in evolutionary changes in embryonic development. This cannot be answered simply by comparing *Hox* genes between two species that differ in some developmental property. A major problem is that species displaying developmental differences may be descended from lineages that separated tens or even hundreds of millions of years ago. Consequently, the original mutations that caused (or contributed to) particular developmental changes are no longer the only differences between the two genomes; even if the original genetic differences still persist, there will be thousands of additional genetic differences between the species.

One way to circumvent this problem is to focus on changes in development that occurred relatively recently, by comparing closely related species with differing modes of development. This approach is being used successfully to investigate the evolution of tail loss and direct development in ascidians (Swalla *et al.*, 1993). Although these studies have not yet implicated *Hox* genes, they do suggest that genes encoding transcription factors have been inactivated during this evolutionary change (Swalla *et al.*, 1993; Jeffery, 1994).

Many of the major developmental changes occurring during chordate evolution, however, occurred hundreds of million years ago; we are forced to tackle them by comparison between very divergent animal species. Despite the problems outlined above, encouraging insights are being made into the role of *Hox* genes in some of these events. Two examples are given: the first dealing with a modification to existing developmental pathways (Burke *et al.*, 1995), the second relating to the invention of new developmental pathways (Garcia-Fernàndez and Holland, 1994).

The study of Burke *et al.* (1995), discussed above in relation to homology, investigated a case in which the coupling between the form of a structure and its position within a segmental series had shifted in evolution. These authors demonstrated that the axial limits of *Hox* gene expression in the developing vertebral column have shifted since the divergence of birds and mammals, in parallel with the ''transposition,'' or shifting correspondence, between vertebral differentiation and segment number. This indicates that evolutionary modification to existing developmental programs has been achieved without change in *Hox* gene number or organization, but was associated with changes in *Hox* gene regulation. Identifying the original mutations that caused transposition of vertebral morphologies is difficult; the data indicate that these mutations either influenced *Hox* gene expression directly or they affected processes upstream from *Hox* gene regulation.

The study by Garcia-Fernàndez and Holland (1994) relates to vertebrate origins. Vertebrates possess morphological characters not seen in the other chordates; these include, but are not restricted to, a suite of characters in the head region. Some have putative homologues in amphioxus or tunicates (e.g., forebrain, hindbrain, possibly the branchial arches, eyes, and otic system; Bone and Ryan, 1978; Holland

et al., 1992; Lacalli *et al.,* 1994); other characters may be truly novel (e.g., cranial ganglia, extensive craniofacial connective tissue, teeth, and bone). The elaboration of primitive chordate characters, plus the origin of vertebrate-specific novelties, must have required modification of developmental programs (Gans and Northcutt, 1983; Holland, 1988, 1992). This in turn may have involved substantial changes in genetic control networks, especially for characters dependent upon differentiation pathways, inductive interactions, or cell behaviors that are also unique to vertebrates (such as the neural crest-derived connective tissue that dominates craniofacial morphogenesis).

The complexity of these developmental changes, plus the antiquity of chordate radiation, suggests that identifying all the genetic mutations that were causative in the evolution of vertebrate development may be impossible. Furthermore, we have no grounds on which to suppose that mutations in *Hox* genes were any more influential than mutations in other developmentally expressed genes (encoding transcription factors, signaling molecules, receptors, etc.). Nonetheless, the demonstration of a single archetypal *Hox* gene cluster in amphioxus (Garcia-Fernàndez and Holland, 1994), in contrast to multiple clusters in all vertebrates (see above), is intriguing. The data reveal that the most extensive phase of elaboration of the chordate *Hox* clusters correlates with the most extensive phase of developmental elaboration and innovation, close to vertebrate origins. This statement becomes more than an isolated correlation when two other lines of evidence are added. First, gene mapping in mammals reveals that the *Hox* gene clusters are part of an extensive ''paralogy group'': a set of physically linked genes, each with linked homologues on two or more chromosomes. The mouse chromosome 6, 11, 15, 2 (or human chromosome 7, 17, 12, 2) paralogy group includes *Hox* gene clusters, collagen genes, retinoic acid receptor genes, *Evx* homeobox genes, glucose transporter genes, actin genes, *GLI/ci*^{D} zinc finger genes, myosin light chain genes, *Wnt* genes, and others (Bentley *et al.,* 1993; Lundin, 1993). This paralogy implies that mutation did not duplicate the *Hox* gene cluster in isolation; rather, it affected an extensive set of linked genes. A second line of evidence for widespread duplication comes from unlinked genes. Within each of the *Cdx, Msx, En, Otx,* and *Emx* class homeobox gene families, plus the *insulin/IGF* family, only single genes have been isolated from amphioxus, whereas jawed vertebrates have two or more homologues (Chan *et al.*, 1990; Holland *et al.*, 1994a,b; N. A. Williams, A. C. Sharman, L. Z. Holland, J.G.F. and P.W.H.H., unpublished data). An exception to this rule is the *Brachyury* gene with at least two very closely related loci in amphioxus (Holland *et al.*, 1995); molecular phylogenetic analyses demonstrate these loci arose by a *Brachyury* gene duplication specific to the cephalochordate lineage.

The *Hox* gene cluster data, the other amphioxus cloning data, and the mammalian linkage data all point to a major phase of gene duplication close to vertebrate origins. This could have involved one or more rounds of tetraploidy of the genome, although this is not proven (Ohno, 1970; Lundin, 1993; Holland *et al.,* 1994b). Cellular and genomic changes that may have allowed major expansion in the number of functional genes are considered by Bird (1995); he suggests that the origin of DNA methylation-based gene repression mechanisms may have permitted this dramatic increase on the vertebrate lineage.

The origin of vertebrates, therefore, was associated with duplication of entire genetic cascades and interacting networks of genes, possibly following the evolution of new modes of methylation. There is no reason to suppose that this directly caused the evolution of new developmental characters; instead, we suggest that the gene duplications overcame a genetic constraint to elaboration of the chordate body plan. After duplication, the gradual modifying effects of mutation, natural selection, genomic flux, and genetic drift would have modified the new genes, gene networks, and gene cascades, inactivating some but eventually co-opting others for the development of new vertebrate characters.

This model for the developmental genetic basis of vertebrate origins makes testable predictions. For example, duplicated genes in vertebrates should display functional redundancy during the development of primitive chordate characters, but not during development of vertebrate-specific features. Similarly, amphioxus and tunicate genes should have functions and expression characteristics in common with their (multiple) vertebrate homologues, but the latter may have additional features associated with the development vertebrate-specific characters only. Preliminary evidence in favor of the latter prediction comes from comparison of vertebrate, amphioxus, and tunicate *Hox* gene expression (Holland, 1992; Holland *et al.,* 1992; Katsuyama *et al.,* 1995; plus this report), but further studies on a range of genes are necessary. Perhaps the most encouraging sign at this stage is that despite the vast evolutionary timescales involved, comparisons of amphioxus, tunicate, and vertebrate *Hox* genes have made a real contribution to understanding the genetic basis for the origins of vertebrate developmental control, in highlighting the importance of gene network duplication.

CONCLUSIONS

In this article we have considered the evolution of chordate homeobox genes belonging to the *Hox* class, only briefly mentioning some of the other homeobox gene classes present in animal genomes (Duboule, 1994). In focusing on *Hox* genes, we do not imply they are the only genes of importance to chordate evolution; however, they do represent the class of developmentally important genes for which the most detailed data have been accumulated from a diversity of chordates. There is every reason to suppose that the study of other genes, in a comparative context, will prove equally informative in due course.

The recent data on *Hox* genes from chordates and other deuterostomes, reviewed here, have relevance to several areas

of developmental and evolutionary biology. Perhaps the most significant insights center around three themes: the molecular evolution of gene families, the developmental genetic basis of homology, and the coevolution of molecules and morphology. Principal conclusions that may be drawn include the following. (1) Multiple *Hox* genes are present in the genomes of enteropneusts, tunicates, amphioxus, hagfish, lampreys, and jawed vertebrates. (2) *Hox* genes are probably organized into gene clusters in all the above taxa; current evidence points to a single cluster in all invertebrate deuterostomes (including amphioxus), but multiple clusters in all vertebrates (including jawless vertebrates). (3) Tandem duplication of *Hox* genes occurred on the deuterostome lineage; for example, duplication of posterior *Hox* genes early in chordate evolution. (4) Some *Hox* genes in tunicates have unusually divergent sequence; some also have introns within the homeobox. (5) The amphioxus *Hox* gene cluster has retained an organization remarkably similar to that inferred for a direct ancestor of vertebrates. (6) Duplication of the *Hox* gene cluster on the vertebrate lineage was followed by loss of some *Hox* genes; gene loss may have occurred on multiple occasions and may be phylogenetically informative. (7) Expression data suggest the primitive role of chordate *Hox* genes was in region-specific patterning of the neural tube; *Hox* genes were coopted for additional roles after *Hox* gene cluster duplication in vertebrates. (8) *Hox* gene expression can reveal homology between vertebrate body plans; morphological similarity and *Hox* gene expression concur in their recognition of homology. (9) *Hox* gene expression can reveal cryptic homology between body plans of divergent chordate subphyla; for example, an extensive region of the neural tube in amphioxus is homologous to the vertebrate hindbrain. This suggests that the vertebrate head evolved by remodeling of an extensive precursor region. (10) *Hox* gene cluster duplication on the vertebrate lineage was part of a major phase of gene or genome duplication; the duplication and cooption of gene networks may have permitted the evolution of vertebrate development and morphology.

There are still some notable gaps in our knowledge concerning *Hox* genes in chordates and other deuterostomes. These include details of the *Hox* gene diversity in pterobranchs, further data from echinoderms (not reviewed here; Dolecki *et al.*, 1988; Pfeffer and von Holt, 1991), and far more extensive *Hox* gene expression and gene linkage data from all relevant taxa. Analyses of lamprey, hagfish, tunicate, and amphioxus *Hox* genes have each raised questions that need addressing, while wider surveys of other gene families must be undertaken to fully understand the amphioxus results in particular. These gaps will undoubtedly be filled over the next few years. If the early results are representative, the comparative analysis of chordate *Hox* genes will make a lasting and significant contribution to our understanding of the relationships between genome evolution, developmental control, and chordate diversity.

ACKNOWLEDGMENTS

We thank many friends and colleagues who have openly discussed their data and ideas and generously allowed us to cite results prior to publication. Particular thanks are due to Drs. A. Adoutte, S. Aparicio, W. Bailey, G. Balavoine, P. Baumgartner, S. Brenner, S. M. Davis, R. Di Lauro, D. Duboule, C. Gans, H. Gee, W. J. Gehring, L. Z. Holland, N. D. Holland, C. Kappen, T. C. Lacalli, L.-G. Lundin, W. McGinnis, K. Peterson, F. H. Ruddle, H. Saiga, A. C. Sharman, S. A. J. Thompson, C. Tomlinson, H. Wada, G. Wagner, N. A. Williams, and G. Wray. The authors' work cited here was supported by the BBSRC, MRC, EMBO, HFSP, The Royal Society, and Acciones Integradas of the British Council/Ministerio de Educación y Ciencia.

REFERENCES

Akam, M. E., Holland, P. W. H., Ingham, P. W., and Wray, G. (Eds.) (1994). The evolution of developmental mechanisms. *Development Suppl.*

Aparicio, S., Morrison, A., Gould, A., Gilthorpe, J., Chaudhuri, C., Rigby, P., Krumlauf, R., and Brenner, S. (1995). Detecting conserved regulatory elements with the model genome of the Japanese puffer fish, *Fugu rubripes*. *Proc. Natl. Acad. Sci. USA* 92, 1684–1688.

Bentley, K. L., Bradshaw, M. S., and Ruddle, F. H. (1993). Physical linkage of the murine Hox-b cluster and nerve growth factor receptor on yeast artificial chromosomes. *Genomics* 18, 43–53.

Bird, A. P. (1995). Gene number, noise reduction and biological complexity. *Trends Genet.* 11, 94–100.

Bone, Q., and Ryan, K. P. (1978). Cupular sense organs in *Ciona* (Tunicata: Ascideacea). *J. Zool.* 186, 417–429.

Brusca, R. C., and Brusca, G. J. (1990). ''Invertebrates.'' Sinauer, Sunderland, MA.

Burke, A. C., Nelson, C. E., Morgan, B. A., and Tabin, C. (1995). *Hox* genes and the evolution of vertebrate axial morphology. *Development* 121, 333–346.

Carrasco, A. E., McGinnis, W., Gehring, W. J., and De Robertis, E. M. (1984). Cloning of a *X. laevis* gene expressed during early embryogenesis coding for a peptide region homologous to *Drosophila* homeotic genes. *Cell* 37, 409–414.

Chan, S. J., Cao, Q.-P., and Steiner, D. F. (1990). Evolution of the insulin superfamily: Cloning of a hybrid insulin/insulin-like growth factor cDNA from amphioxus. *Proc. Natl. Acad. Sci. USA* 87, 9319–9323.

Di Gregorio, A., Spagnuolo, A., Ristoratore, F., Pischetola, M., Aniello, F., Branno, M., Cariello, L., and Di Lauro, R. (1995). Cloning of ascidian homeobox genes provides evidence for a primordial chordate cluster. *Gene* 156, 253–257.

Dickinson, W. J. (1995). Molecules and morphology: Where's the homology? *Trends Genet.* 11, 119–121.

Dolecki, G. J., Wang, G., and Humphreys, T. (1988). Stage- and tissue-specific expression of two homeobox genes in sea urchin embryos and adults. *Nucl. Acids Res.* 16, 11543–11558.

Duboule, D. (1994). ''Guidebook to the Homeobox Genes.'' Oxford Univ. Press, Oxford.

Duboule, D., and Dollé, P. (1989). The structural and functional organization of the murine *Hox* gene family resembles that of *Drosophila*. *EMBO J.* 8, 1497–1505.

Gans, C., and Northcutt, R. G. (1983). Neural crest and the origin of the vertebrates: A new head. *Science* 220, 268–274.

Garcia-Fernàndez, J., and Holland, P. W. H. (1994). Archetypal organization of the amphioxus *Hox* gene cluster. *Nature* 370, 563–566.

Gaunt, S. J. (1991). Expression patterns of mouse *Hox* genes: Clues

to an understanding of developmental and evolutionary strategies. *BioEssays* 13, 505–513.

Ge, T., Lee, H. M., and Tomlinson, C. R. (1994). Identification of an *antennapedia*-like homeobox gene in the ascidians *Styela clava* and *S. plicata*. *Gene* 147, 219–222.

Graham, A., Papalopulu, N., and Krumlauf, R. (1989). The murine and *Drosophila* homeobox gene complexes have common features of organization and expression. *Cell* 57, 367–378.

Halanych, K. M. (1995). The phylogenetic position of the pterobranch hemichordates based on 18S rDNA sequence data. *Mol. Phylogenet. Evol.* 4, 72–76.

Hall, B. K. (1992). ''Evolutionary Developmental Biology.'' Chapman & Hall, London.

Hall, B. K. (Ed.) (1994). ''Homology: The Hierarchical Basis of Comparative Biology.'' Academic Press, San Diego.

Holland, P. W. H. (1988). Homeobox genes and the vertebrate head. *Development* 103, (*Suppl*), 17–24.

Holland, P. W. H. (1992). Homeobox genes in vertebrate evolution. *BioEssays* 14, 267–273.

Holland, P. W. H., and Graham, A. (1995). Evolution of regional identity in the vertebrate nervous system. *Perspect. Dev. Neurobiol.* 3, 17–27.

Holland, N. D., and Holland, L. Z. (1989). Fine structural study of the cortical reaction and formation of the egg coats in a lancelet (=amphioxus), *Branchiostoma floridae* (Phylum Chordata: Subphylum Cephalochordata = Acrania). *Biol. Bull. Mar. Biol. Lab. Woods Hole* 176, 111–122.

Holland, N. D., and Holland, L. Z. (1993). Embryos and larvae of invertebrate deuterostomes. *In* ''Essential Developmental Biology: A Practical Approach'' (C. D. Stern and P. W. H. Holland, Eds.), pp. 21–32. IRL Press at Oxford Univ. Press, Oxford.

Holland, P. W. H., Holland, L. Z., Williams, N. A., and Holland, N. D. (1992). An amphioxus homeobox gene: Sequence conservation, spatial expression during development and insights into vertebrate evolution. *Development* 116, 653–661.

Holland, P. W. H., Garcia-Fernàndez, J., Holland, L. Z., Williams, N. A., and Holland, N. D. (1994a). The molecular control of spatial patterning in amphioxus. *J. Mar. Biol. Assoc. UK* 74, 49–60.

Holland, P. W. H., Garcia-Fernàndez, J., Williams, N. A., and Sidow, A. (1994b). Gene duplications and the origins of vertebrate development. *Development Suppl.* 125–133.

Holland, P. W. H., Korschorz, B., Holland, L. Z., and Herrmann, B. G. (1995). Conservation of *Brachyury* (*T*) genes in amphioxus and vertebrates: developmental and evolutionary implications. *Development*, in press.

Hunt, P., Whiting, J., Nonchev, S., Sham, M.-H., Marshall, H., Graham, A., Cook, M., Allemann, R., Rigby, P. W. J., Gulisamo, M., Faiella, A., Boncinelli, E., and Krumlauf, R. (1991). The branchial *Hox* code and its implications for gene regulation, patterning of the nervous system and head evolution. *Development Suppl.* 2, 63–77.

Janvier, P. (1993). Patterns of diversity in the skull of jawless fishes. *In* ''The Skull, Vol. 2'' (J. Hanken and B. K. Hall, Eds.), pp. 131–188. Univ. of Chicago Press, Chicago.

Jefferies, R. P. S. (1986). ''The Ancestry of the Vertebrates.'' Br. Museum (Natural History), London.

Jeffery, W. R. (1994). A model for ascidian development and developmental modifications during evolution. *J. Mar. Biol. Assoc. UK* 74, 35–48.

Kappen, C., and Ruddle, F. H. (1993). Evolution of a regulatory gene family: *HOM/Hox* genes. *Curr. Opin. Genet. Dev.* 3, 931–938.

Kappen, C., Schughart, K., and Ruddle, F. H. (1989). Two steps in the evolution of Antennapedia-class vertebrate homeobox genes. *Proc. Natl. Acad. Sci. USA* 86, 5459–5463.

Kappen, C., Schughart, K., and Ruddle, F. H. (1993). Early evolutionary origin of major homeodomain sequence classes. *Genomics* 18, 54–70.

Katsuyama, Y., Wada, S., Yasugi, S., and Saiga, H. (1995). Expression of the *labial* group gene *HrHox-1* and its alteration by retinoic acid in development of the ascidian *Halocynthia roretzi*. *Development* 121, 3197–3205.

Kenyon, C., and Wang, B. (1991). A cluster of Antennapedia class homeobox genes in a nonsegmented animal. *Science* 253, 516–517.

Kido, Y., Aono, M., Yamaki, T., Matsumoto, K.-I., Murata, S., Saneyoshi, M., and Okada, N. (1991). Shaping and reshaping of salmonid genomes by amplification of tRNA-derived retroposons during evolution. *Proc. Natl. Acad. Sci. USA* 88, 2326–2330.

Lacalli, T. C., Holland, N. D., and West, J. E. (1994). Landmarks in the anterior central nervous system of amphioxus larvae. *Phil. Trans. Roy. Soc. B* 344, 165–185.

Lundin, L. (1993). Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. *Genomics* 16, 1–19.

Gaskell, W. H., MacBride, E. W., Starling, E. H., Goodrich, E. S., Gadow, H., Smith Woodward, A., Dendy, A., Lankester, E. R., Chalmers Mitchell, P., Gardiner, J. S., Stebbing, T. R. R., and Scott, D. H. (1909). Discussion on the origin of vertebrates. *Proc. Linn. Soc. London* 122.

McGinnis, W., and Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell* 68, 283–302.

McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A., and Gehring, W. J. (1984). A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoa. *Cell* 37, 403–408.

Minelli, A. (1993). ''Biological Systematics: The State of the Art.'' Chapman & Hall, London.

Misof, B. Y., and Wagner, G. P. (1996). Evidence for four Hox clusters in the teleost, *Fundulus heteroclitus*. *Mol. Phylogenet. Evol.*, in press.

Molven, A., Hordvik, I., Njolstad, P. R., Van Ghelue, M., and Fjose, A. (1992). The zebrafish homeobox gene *hox*[*zf-114*]: Primary structure, expression pattern and evolutionary aspects. *Int. J. Dev. Biol.* 36, 229–237.

Murphy, P., and Hill, R. E. (1991). Expression of the mouse *labial*-like homeobox-containing genes, *Hox 2.9* and *Hox 1.6*, during segmentation of the hindbrain. *Development* 111, 61–74.

Ohno, S. (1970). ''Evolution by Gene Duplication.'' Springer-Verlag, Heidelberg.

Papalopulu, N., Lovell-Badge, R., and Krumlauf, R. (1991). The expression of murine *Hox-2* genes is dependent on the differentiation pathway and displays a colinear sensitivity to retinoic acid in F9 cells and Xenopus embryos. *Nucleic Acids Res.* 19, 5497–5506.

Pavell, A. M., and Stellwag, E. J. (1995). Survey of *Hox*-like genes in the teleost *Morone saxatilis:* Implications for evolution of the *Hox* gene family. *Mar. Mol. Biol. Biotech.*, in press.

Pendleton, J. W., Nagai, B. K., Murtha, M. T., and Ruddle, F. H. (1993). Expansion of the *Hox* gene family and the evolution of chordates. *Proc. Natl. Acad. Sci USA* 90, 6300–6304.

Peterson, K. J. (1995). A phylogenetic test of the calcichordate scenario. *Lethaia* 28, 25–38.

Pfeffer, V. L., and von Holt, C. (1991). Stage-specific and adult tis-

sue-specific expression of a homeobox gene in embryo and adult *Parechinus angulosus* sea urchins. *Gene* 108, 219–226.

Philippe, H., Chenuil, A., and Adoutte, A. (1994). Can the Cambrian explosion be inferred through molecular phylogeny? *Development Suppl.* 15–25.

Prince, V., and Lumsden, A. (1994). *Hoxa-2* expression in normal and transposed rhombomeres: Independent regulation in the neural tube and neural crest. *Development* 120, 911–923.

Raff, R. A., and Kaufman, T. C. (1983). ''Embryos, Genes and Evolution.'' MacMillan Co., New York.

Reeck, G. R. C., De Haan, C., Doolittle, D. C. *et al.* (1987). Homology in proteins and nucleic acids: A terminology muddle and a way out of it. *Cell* 50, 667.

Roth, V. L. (1984). On homology. *Biol. J. Linn. Soc.* 22, 13–29.

Ruddle, F. H., Bentley, K. L., Murtha, M. T., and Risch, N. (1994). Gene loss and gain in the evolution of the vertebrates. *Development Suppl.* 155–161.

Saiga, H., Mizokami, A., Makabe, K. W., Satoh, N., and Mita, T. (1991). Molecular cloning and expression of a novel homeobox gene *AHox1* of the ascidian *Halocynthia roretzi*. *Development* 111, 821–828.

Sankoff, D., Leduc, G., Anotine, N., Paguin, B., Lang, B. F., and Cedergren, R. (1992). Gene order comparisons for phylogenetic inference: Evolution of the mitochondrial genome. *Proc. Natl. Acad. Sci. USA* 89, 6575–6579.

Satoh, N. (1994). ''Developmental Biology of Ascidians.'' Cambridge Univ. Press, Cambridge, UK.

Schubert, F. R., Nieselt-Struwe, K., and Gruss, P. (1993). The *Antennapedia*-type homeobox genes have evolved from three precursors separated early in metazoan evolution. *Proc. Natl. Acad. Sci. USA* 90, 143–147.

Scott, M. P. (1992). Vertebrate homeobox gene nomenclature. *Cell* 71, 551–553.

Simeone, A., Acampora, D., Arcioni, L., Andrews, P. W., Boncinelli, E., and Mavilio, F. (1990). Sequential activation of HOX2 homeobox genes by retinoic acid in human embryonal carcinoma cells. *Nature* 346, 763–766.

Slack, J. M. W., Holland, P. W. H., and Graham, C. F. (1993). The zootype and the phylotypic stage. *Nature* 361, 490–492.

Sordino, P., van der Hoeven, F., and Duboule, D. (1995). *Hox* gene expression in teleost fins and the evolution of vertebrate digits. *Nature* 375, 678–681.

Stock, D. W., and Whitt, G. S. (1992). Evidence from 18S ribosomal RNA sequences that lampreys and hagfish form a natural group. *Science* 257, 787–789.

Swalla, B. J., Makabe, K. W., Satoh, N., and Jeffery, W. R. (1993). Novel genes expressed differentially in ascidians with alternate modes of development. *Development* 119, 307–318.

Wada, H., and Satoh, N. (1994). Details of the evolutionary history from invertebrates, as deduced from the sequences of 18S rDNA. *Proc. Natl. Acad. Sci. USA* 91, 1801–1804.

Willey, A. (1894). ''Amphioxus and the Ancestry of the Vertebrates.'' MacMillan Co., New York.

Wilkinson, D. G., Bhatt, S., Cook, M., Boncinelli, E., and Krumlauf, R. (1989). Segmental expression of *Hox 2* homeobox genes in the developing hindbrain. *Nature* 341, 405–409.

Received for publication October 24, 1995
Accepted October 24, 1995